Identifying stimuli of different perceptual categories in mixed blocks of trials: evidence for cost in switching between computational processes.

Los, S.A.

published in
Journal of Experimental Psychology: Human Perception and Performance
1999

DOI (link to publisher)
10.1037/0096-1523.25.1.3

document version
Publisher's PDF, also known as Version of record

Journal of Experimental Psychology:
Human Perception and Performance
1999, Vol. 25, No. 1, 3–23

0096-1523/99/$3.00

Identifying Stimuli of Different Perceptual Categories in Mixed Blocks of Trials: Evidence for Cost in Switching Between Computational Processes

Sander A. Los
Vrije Universiteit

Responding to stimuli of different perceptual categories is usually faster when the categories are presented isolated from each other, in pure blocks, than when they are presented intermixed, in mixed blocks. According to criterion models, these perceptual mixing costs result from the use of a less conservative response criterion in pure than in mixed blocks. According to alternate processing models, mixing costs result from time-consuming switching in mixed blocks between different computational processes called on by the different perceptual categories. In 5 experiments, participants had to identify number stimuli of different categories. The results showed clear mixing costs whenever these categories differed in their assumed computational processing requirements but not when they differed on features that seemed trivial from a computational viewpoint. The results favor the alternate processing conception.

In the pure/mixed design, the levels of an independent variable are presented both in isolation, in pure blocks, and intermixed across trials, in mixed blocks. The common finding is that people respond slower to either level of the variable in mixed blocks than in pure blocks, a difference referred to as *mixing costs.* Mixing costs have been reported for a large number of variables, including stimulus intensity, stimulus quality, stimulus–response compatibility, foreperiod, and task set (see Los, 1996, for a review). In this article I focus on perceptual mixing costs, which are found when a perceptual variable is presented in the pure/mixed design. For instance, Van Duren and Sanders (1988) and Los (1994) had participants identify intact and degraded digits (i.e., two levels of the variable "stimulus quality") both in pure and mixed blocks of trials. In those studies, both intact and degraded digits showed mixing costs in that they were more slowly identified in mixed blocks than in pure blocks.

Mixing costs have recently evoked interest among psychologists because they provide insights into human control of action. I discuss two issues: (a) the source of mixing costs and (b) the mechanism underlying mixing costs. This article is mainly concerned with the second issue, but for the sake of completeness I begin by addressing the first issue in some detail.

The Source of Mixing Costs

According to the *endogenous view,* mixing costs reflect the fact that pure blocks enable the participant to prepare better for forthcoming events than do mixed blocks. This view developed from studies that used mixed blocks in which the levels of the variable were randomized across trials (hereinafter referred to as "randomly mixed blocks"). In these blocks, participants cannot predict which level will be presented on a trial. The uncertainty thus induced is assumed to restrain participants from effectively preparing for forthcoming events and accordingly to reduce their performance relative to that in pure blocks. In this view, the participant's preparatory activity proceeds from an internal source, possibly a central executive (Baddeley, 1986) or a supervisory attentional system (Norman & Shallice, 1986), and so the theoretically important implication is that mixing costs are endogenous in nature. Consequently, mechanisms accounting for mixing costs on the basis of uncertainty indicate where and how endogenous preparation enhances the anticipated flow of information (e.g., Grice, 1968; Stoffels, 1996; Van Duren & Sanders, 1988).

The endogenous view used to be dominant in the literature of mixing costs; witness the common use of variable labels such as "stimulus uncertainty" or "event uncertainty" (Grice, 1968; Sperling & Dosher, 1986; Van Duren & Sanders, 1988), with the levels "low" and "high," instead of the more neutral label of "block type," with the levels "pure" and "mixed," respectively. This is unfortunate because block type manipulates more than uncertainty per se. This is perhaps most notable in studies on task switching, in which participants carry out two tasks, isolated from each other (i.e., in pure blocks) as well as in a perfectly predictable order across trials of mixed blocks. In spite of the resulting equal predictability of events in pure and mixed blocks, researchers using this design have reported mixing costs almost without exception (e.g., Allport, Styles, & Hsieh, 1994; Jersild, 1927; Meiran, 1996; Rogers & Monsell, 1995; Spector & Biederman, 1976). The results of those studies suggest that the mere difference in intertrial variabil-

I am grateful to Andries Sanders, Gordon Baylis, Rob van der Lubbe, and an anonymous reviewer for helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Sander A. Los, Department of Psychonomics, Vrije Universiteit, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands. Electronic mail may be sent to sander@psy.vu.nl.

ity between pure and mixed blocks is a sufficient condition for mixing costs to occur. Consequently, within the task-switching tradition, mixing costs are usually attributed to the need of the mental system to adjust to different demands when the level of the variable alternates on subsequent trials. Such "switch costs" necessarily translate into mixing costs because switching between the levels of a variable occurs only in mixed blocks, but not in pure blocks. The theoretically important implication of this view is that mixing costs are *exogenous* in nature, as they are assumed to reflect the degree to which current mental activities are supported or impeded by preceding ones.

Even in the case of an equally low uncertainty in mixed blocks as in pure blocks (as realized in the task-switching paradigm), the finding of mixing costs does not provide unequivocal support for the exogenous view, however. Alternatively, mixing costs could reflect block differences in *mental load,* in the sense that participants must maintain readiness for either level of the variable in mixed blocks, but only for one level in pure blocks. As a result, less capacity is available for processes called on in mixed blocks than for these processes in pure blocks, thus reducing processing efficiency. The key feature of this conception of mental load is its static character, because it assumes that the participant continuously maintains preparedness for all possible events that may occur in a given block of trials. Therefore, even when participants know on each trial of a mixed block which level of the independent variable is forthcoming, mixing costs are still predicted by a mental load account because readiness for the alternative level of the variable should be maintained for the benefit of later trials (e.g., Ilan & Miller, 1994).

The list given so far may not be exhaustive, but it captures the main theoretical positions regarding the source of mixing costs. In a previous review (Los, 1996) I examined the source of mixing costs for several variables that load stages varying from early perceptual processing to late motor adjustment, as revealed by additive factors research (e.g., Sanders, 1990). Regarding the perceptual level, I discussed two lines of evidence indicating that mixing costs occurring here are predominantly exogenous in nature. The first line derives from studies that deconfounded uncertainty and intertrial variability. These studies have used mixed blocks that combine the high intertrial variability of randomly mixed blocks and the low uncertainty of pure blocks. This is achieved either by presenting the stimulus categories (i.e., the levels of a perceptual variable) in a fixed order (e.g., AABBAABB, etc.), such as in studies on task switching, or by informing the participant in advance of each trial of the forthcoming stimulus category. For several perceptual variables, it has been shown that performance in these blocks was still clearly poorer than that in pure blocks and not much different from that in randomly mixed blocks (Los, 1994, 1997; Maljkovic & Nakayama, 1994; Sanocki, 1988). Thus, the level of uncertainty per se appears to be a poor predictor of perceptual mixing costs, which argues strongly against an endogenous source. A second line of evidence derives from the study of sequential effects in mixed blocks, in which performance is separately analyzed for category repetition

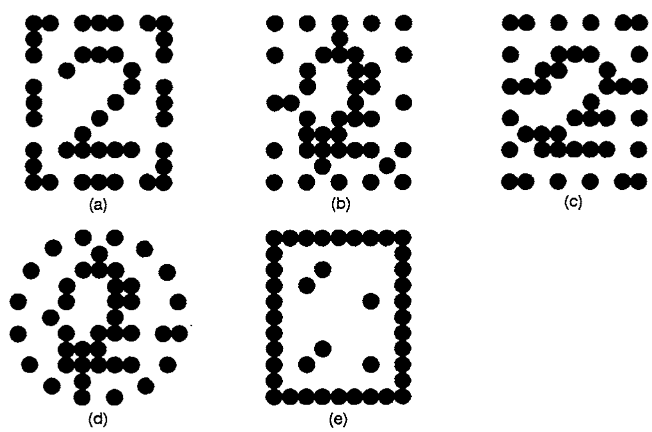

Figure 1. Examples of the stimulus categories, depicted as negatives, occurring in Experiments 1–4. All examples represent the digit 2: the intact category (a), the noise-degraded category with 12 noise dots within a rectangular frame (b), the noise-degraded category with 8 noise dots within a rectangular frame (c), the noise-degraded category with 12 noise dots within a round frame (d), and the segment-deleted category (e). All stimuli are composed of an equal number of dots.

trials (i.e., trials on which the stimulus category of the preceding trial is repeated) and category alternation trials (i.e., trials on which the stimulus category is different from that of the preceding trial). Results of these studies have revealed that perceptual mixing costs are relatively large on alternation trials and relatively small or even absent on repetition trials (Los, 1997; Maljkovic & Nakayama, 1994). The dynamic character of these sequential effects is particularly hard to reconcile with a mental load account, which conceives mixing costs as resulting from a difference in static load between pure and mixed blocks.[1]

Elsewhere (Los, 1997), I confirmed these results for the stimulus categories used in the present research (see Figure 1). Together, the results provide strong support for the exogenous nature of perceptual mixing costs. According to this view, mixing costs are a derivative of switch costs, which are autonomous and escape the influence of an endogenous preparatory mechanism.

Mechanisms Explaining Mixing Costs

Although clearly constraining possible explanatory mechanisms, the finding that mixing costs have an exogenous source is not diagnostic about the components of which a mechanism should consist. As will become apparent later,

[1] Note that this finding of sequential effects is also at least slightly problematic for the endogenous view. To account for this finding, this view has to assume that participants adopt a *strategy* of using the event that occurred on the last trial as a point of reference for current preparation. Although this view has been defended in the literature (e.g., Kirby, 1980), empirical data generally support the opposite view that in a two-choice task, people are inclined to prepare for an alternation of events (i.e., the gambler's fallacy; Jarvik, 1951; Soetens, Boer, & Hueting, 1985).

the same components that are proposed to operate under endogenous control in one mechanism may serve under exogenous control in another mechanism and vice versa. Therefore, to answer the question of *what* is adjusted on alternation trials of mixed blocks, one should obtain additional information.

One relevant source of information is that mixing costs are typically larger for the level of the variable that is responded to faster in pure blocks. For instance, Van Duren and Sanders (1988) and Los (1994) had participants identify intact and degraded digits in pure and mixed blocks of trials and found that mixing costs were greater for intact stimuli (the "fast" level) than for degraded stimuli (the "slow" level). This pattern of *asymmetric* mixing costs is quite general, because it holds for a wide range of variables (see Los, 1996, for a review).

The finding of asymmetric mixing costs has been the impetus for an important class of models referred to here as "criterion models" (e.g., Grice, 1968; Lupker, Brown, & Colombo, 1997; Strayer & Kramer, 1994a, 1994b). Basically, these models postulate two components: an input function and a response criterion. The input function describes how information about the correct response gradually accumulates in time upon the presentation of the imperative stimulus. As soon as the input function exceeds the independently set response criterion, the participant emits a response. Within this general framework, criterion models may show considerable variations, mainly concerning the form of the input function and whether the criterion is set in the information domain or in the time domain. When the criterion is set in the information domain, participants emit a response as soon as a fixed amount of information has accumulated (e.g., Grice, 1968; McClelland, 1979; Ratcliff, 1978; Strayer & Kramer, 1994a, 1994b); when set in the time domain, the criterion has the characteristic of a deadline, causing participants to emit a best guess when processing has not finished when the deadline expires (e.g., Lupker et al., 1997; Ollman, 1966; Ollman & Billington, 1972; Ratcliff, 1978). Another issue that differentiates among criterion models is how the setting of the criterion is controlled. In addition to the conventional idea that endogenous factors, such as motivation and compliance with instructions, determine the setting of the criterion, some models assume that the fine-tuning of the criterion is regulated by recent experiences (e.g., Strayer & Kramer, 1994a, 1994b; Treisman & Williams, 1984). This suggests the possibility that the criterion is exogenously tuned, which precludes the rejection of criterion models on the basis of the finding that mixing costs are exogenous in origin.

Disregarding possible variations, different criterion models offer an essentially similar account of asymmetric mixing costs. The central idea is that the response criterion is differentially set in pure and mixed blocks of trials. For instance, Lupker et al. (1997, Experiment 4) had participants quickly read aloud words of a regular or an irregular orthography, either in pure or in mixed blocks of trials. For the regular words, they found mixing costs, in that these words were named faster in pure blocks than in mixed blocks. For the irregular words, they reported mixing benefits, in that these words were named faster in mixed blocks than in pure blocks. To account for these results, Lupker et al. assumed that pure blocks enable the participant to optimally set the position of a processing deadline, because of relatively homogeneous finishing times of fast identification processes for regular words on the one hand and of slow identification processes for irregular words on the other. In mixed blocks, however, finishing times are a mixture of these fast and slow finishing processes, and so it is not clear where the deadline is optimally set. On the basis of their data, Lupker et al. argued that their participants used a deadline in between the deadlines used in pure blocks for regular and irregular words. As a result, regular words showed mixing costs, because of a relatively sloppy deadline in mixed blocks, whereas irregular words showed mixing benefits, because of a relatively sharp deadline in mixed blocks. In support of this view, Lupker et al. found that accuracy differences across their conditions tended to conform to the predictions of the speed–accuracy trade-off.

An alternative class of models that may account for asymmetric mixing costs is referred to here as "alternate processing models." Central to these models is the assumption that different levels of a variable call on different processing (e.g., Allport et al., 1994; Stoffels, 1996; Sudevan & Taylor, 1987; Van Duren & Sanders, 1988). Again, considerable variations may be found among different models of this class. One issue is the grain size of what is understood by different processing. A large grain size is embodied in models that distinguish only two different types of processing, usually referred to as "processing routes" (e.g., Monsell, Patterson, Graham, Huges, & Milroy, 1992; Paap, Williams, & Johansen, 1992; Stoffels, 1996; Van Duren & Sanders, 1988). The grain size is much smaller in models proposing that highly differentiated processing derives from the specific adjustment of parameters (e.g., Sanocki, 1987, 1988). Another issue is how the different processes are controlled, analogous to the nature of criterion control in criterion models. Some models focus exclusively on endogenous control and assume that the participant may emphasize and deemphasize processes in preparation for future actions (e.g., Stoffels, 1996; Van Duren & Sanders, 1988). Other models favor the role of exogenous control instead and assume that the mental apparatus is inert, which implies that processing efficiency depends on whether relevant processes have been activated shortly before (e.g., Allport et al., 1994; Los, 1994).

An illustration of an alternate processing account is taken from the realm of task switching, where a consensus exists that different task levels involve different processing. Allport et al. (1994, Experiment 5) used a Stroop task, in which two exemplars (Stimulus 1 [S1] and Stimulus 2 [S2]) of incongruent color words (e.g., the word *red* printed in blue) were sequentially presented on each trial. Participants had to name aloud either the word (i.e., word reading) or the color in which the word was printed (i.e., color naming). In pure blocks, participants had to execute the same task with respect to S1 and S2 (i.e., either word reading or color naming), whereas in mixed blocks they had to execute both tasks with respect to S1 and S2 in a fixed order (i.e., either word reading and color naming or color naming and word reading). Regarding the time to respond to S2, the results showed the typical pattern of asymmetric mixing costs: namely, sizable mixing costs for the fast word-reading task but no mixing costs for the slow color-naming task. Similar results were found for error proportions (PEs): that is, mixing costs for the fast word-reading task but no costs for the slow color-naming task.

In view of the absence of any trade-off of speed for accuracy, these data are not readily accounted for in terms of criterion shifts between pure and mixed blocks. Instead, it is more likely that the control of word-reading and color-naming processes causes greater problems in mixed blocks than in pure blocks. Thus, from the perspective of endogenous control, it could be suggested that pure blocks allow for better preparation for word-reading *or* color-naming processes than mixed blocks allow for word-reading *and* color-naming processes. In this view, the asymmetry of mixing costs is typically accounted for in terms of a "worst-case scenario" in mixed blocks, implying that participants prepare optimally for the more difficult, or slower, event at the expense of the more easy, or fast, event (e.g., Monsell et al., 1992; Stoffels, 1996; Van Duren & Sanders, 1988). Allport et al. (1994) rejected an account based on endogenous control, however, because they found that their mixing costs hardly decreased as the response stimulus interval between S1 and S2 increased from 20 to 1,100 ms, suggesting that participants were unable to exploit any extra time during this interval to enhance their preparatory state for the anticipated forthcoming task. Instead, they proposed that a task set is subject to inertia, which facilitates its reuse and hampers the shift to an alternative set. To explain the asymmetry of mixing costs, they assumed that inertia is stronger as a task set demands more control. In theoretical contributions to the Stroop task it is commonly assumed that color naming is control demanding but that word reading proceeds automatically. Thus, the mental system must overcome more inertia in switching from color naming to word reading than vice versa, leading to greater mixing costs for word reading than for color naming, as observed.

The results reported by Lupker et al. (1997) on the one hand and by Allport et al. (1994) on the other clearly demonstrate that both criterion shifts and alternate processing may contribute to mixing costs. Which one of these is dominant depends on the variable under investigation. In the perceptual domain, this matter awaits further examination. The major objective of this article, then, is to provide evidence that supports either the criterion-based account or the alternate processing account of perceptual mixing costs.

The Experimental Approach

The straightforward approach to distinguishing between a criterion account and an alternate processing account of mixing costs is to examine the effects of a speed–accuracy trade-off. Ideally, a criterion account predicts that any mixing costs for reaction time (RT) are compensated for by mixing benefits for PEs and vice versa, whereas alternate processing accounts do not predict the effects of a speed–accuracy trade-off. This approach might be effective when mixing costs are relatively sizable, so that there is room for differences in PEs to become manifest. For perceptual variables, however, mixing costs are typically small, on the order of 10–30 ms, and it is questionable whether reliable effects of a speed–accuracy trade-off could be expected. For instance, Colombo and Tabossi (1992, Experiment 2) used a deadline method to vary speed stress, but they failed to find any accuracy decrement compensating for a 60-ms speedup of RT (see also Lupker et al., 1997, on this point).

In this article I take an alternative approach and explore the stimulus conditions under which mixing costs occur. For this purpose, several stimulus categories were created and pairwise tested in the pure/mixed design. This enabled the analysis of mixing costs as a function of variation along the stimulus dimensions of interest. As I show, this analysis reveals a great deal about possible mechanisms underlying mixing costs.

General Method

Participants

Participants were male and female students from the Vrije Universiteit with normal or corrected-to-normal vision. They were paid for their services. Each participant took part in only one experiment.

Apparatus

An Olivetti PCS microcomputer equipped with a 386 SX processor controlled the experiment and recorded the data. The stimuli were presented on an Olivetti video graphics array screen. The software package ERTS was used for composing the experimental task, whereas the EXKEY interface between the computer and an external response panel allowed RTs to be measured to the nearest millisecond (Beringer, 1992). The response panel contained a 22 × 27 cm, 10° tilted plane on which four microswitches were mounted, each covered by a round response key 2.5 cm in diameter. The response keys were arranged as follows. Two outward upper keys were horizontally aligned, 4 cm from the top of the plane and at a distance of 13 cm from each other. In between these keys, two inward lower keys were horizontally aligned, 6.5 cm from the top of the plane and at a distance of 4.5 cm from each other.

Task

Participants were seated at a distance of about 50 cm from the screen with their middle fingers resting on the outward, upper response keys and their index fingers on the inward, lower response keys. A trial started with the appearance of a visual warning signal at the center of the screen for 1,000 ms. This signal was a white plus sign, 2.5 × 2.5 cm, subtending a visual angle of 2.9° × 2.9° given a viewing distance of 50 cm. The luminance of the signal was 7.0 cd/m^2 against a background of 0.5 cd/m^2. A blank interval of 300 ms separated the offset of the warning signal from the presentation of the imperative stimulus at the same location. The imperative stimulus represented a numerical value from the set {2, 3, 4, 5}. The various representations of these values are described in detail below. Participants responded by pressing the appropriate response key. The assignment of the numerical values to the response keys was as follows: 2, left upper key; 3, left lower key; 4, right lower key; and 5, right upper key. The imperative stimulus disappeared as soon as the participants responded or when a maximum time of 1,500 ms had expired. A blank interval of 200 ms separated the moment of the offset of the imperative stimulus from the onset of the warning signal of the next trial.

Stimuli

All imperative stimuli used in the first four experiments of the present study were digits, representing the numerical values 2, 3, 4, and 5 (see Figure 1). The size of the digits was 3.5 × 2.5 cm, and they subtended a visual angle of 4.0° × 2.9° given a viewing distance of 50 cm. The digits were surrounded either by a rectangular frame of 5.5 × 4.5 cm, subtending a visual angle of 6.3° × 5.2°, or by a round frame 5.7 cm in diameter, subtending a visual angle of 6.5°. The digits and their surrounding frames were composed of 42 white dots 0.5 mm in diameter with a luminance of 7.0 cd/m^2 against a black background of 0.5 cd/m^2. Figure 1 shows an example of the digit 2 for each stimulus category used in Experiments 1–4. The main categories comprised *intact stimuli* (Figure 1a), *noise-degraded stimuli* (Figures 1b, 1c, and 1d), and *segment-deleted stimuli* (Figure 1e). Intact stimuli contained a digit composed of 14 dots and a surrounding rectangular frame of 28 dots. The dots in the frame were adjacent, except for eight 0.5-mm openings, two at each side of the frame (cf. Figure 1a). Noise-degraded and segment-deleted digits were derived from the intact primitives. In the noise-degraded condition, dots were symmetrically deleted from the frame and pseudorandomly dispersed around the digit. The number of noise dots was either 12 (Figure 1b) or 8 (Figure 1c). In one condition, digits degraded by 12 noise dots were surrounded by the round frame composed of 16 regularly distributed dots (Figure 1d). Finally, in the segment-deleted condition, 8 dots were pseudorandomly deleted from the intact digit and inserted in the open spaces of the frame, yielding a complete frame (Figure 1e).

There were four different instances of each noise-degraded and segment-deleted digit, constituted by different noise patterns and segment deletions, respectively. Consequently, each noise-degraded and segment-deleted category (i.e., Figures 1b–1e) comprised a set of 16 instances, constituted of four instances of each of four digits. The category of intact digits (i.e., Figure 1a) comprised a set of only four instances, one of each digit. The 12- and 8-dot noise patterns in the noise-degraded condition (Figures 1b and 1c) did not have any systematic overlap, but the noise patterns of the 12-dot noise conditions that were surrounded by the rectangular frame and the round frame (Figures 1b and 1d) were almost identical. In the latter case, only the noise dots that occurred at coordinates within the rectangular frame that had no analog within the round frame (i.e., at and near the corners of the rectangular frame) were transferred to coordinates where this was reversed (i.e., to the left and right sides of the circle). Furthermore, the noise-degraded and segment-deleted stimuli were composed in such a way that the digits remained unambiguously identifiable. For the noise-degraded stimuli, this meant that no conspicuous alternative digit emerged from a combination of the digit at issue and some dots of the noise pattern. For the segment-deleted stimuli, this meant that the distinguishing features of each digit remained relatively unaffected.

Block Type

Blocks of trials varied in the proportion of occurrence of either of two stimulus categories. In pure blocks, only one category occurred. In mixed blocks, the following proportions were used (a) 1:1, indicating an equal probability of occurrence of either category on each trial and (b) 1:3/3:1, indicating a 75% probability of occurrence of one category, the "frequent" category, and a 25% probability of occurrence of the other, "infrequent" category on each trial.

Each task-relevant alternative (i.e., the value of the digit) had an equal (.25) probability of being presented on a trial across all blocks and experiments. To ensure that each (task-irrelevant) instance of each digit also had an equal probability of being presented across blocks, I used a balancing procedure in Experiments 2–4. In these experiments, categories of degraded stimuli were used, of which each digit had four different instances. This enabled the creation of a set of 16 instances for each block, from which the imperative stimuli were randomly drawn. In a pure block, this set contained all 16 instances of the category at issue. In a mixed-1:1 block, eight instances of either category were included in the set, two instances of each digit. The remaining eight instances of either category occurred in a complementary mixed-1:1 block, presented at another moment in the experiment. An analogous procedure was followed for mixed-3:1 blocks. Here a set was created by including four instances of the infrequent category, one instance of each digit, and excluding four instances of the frequent category, one instance of each digit. This set was adjusted for subsequent mixed-3:1 blocks, such that across four blocks, each instance of the infrequent category was included precisely once and each instance of the frequent category was excluded precisely once. Finally, for each mixed-3:1 block, a mixed-1:3 counterpart was created containing all complementary instances.

Design and Procedure

Stimulus category and block type were factorially combined as independent variables, and the resulting conditions were counterbalanced across participants. Individual RTs and PEs were the dependent variables.

Participants came to the laboratory for about 2 hr. As a first introduction to the stimuli, all instances of the degraded digits occurring in an experiment were presented on the screen in random order. Participants had to name aloud the values these stimuli represented, and they were corrected by the experimenter if necessary. This acquaintance procedure continued until all alternatives could be correctly named. In the task instruction, speed and accuracy were equally emphasized. Each block of trials was preceded by a text on the screen for 10 s that informed participants which stimulus categories would occur in that block and in what proportion. This block information was replaced by a text for 5 s that summoned participants to position their fingers on the response keys, whereupon a block started. Participants received one or two blocks of practice in each condition depending on the stability of their performance. During practice, participants received feedback on RTs and PEs after each block of trials. After practice, there were four (Experiments 1–4) or two (Experiment 5) consecutive experimental sessions. A session comprised all pure and mixed blocks, which were presented in alternation as much as possible. The number of blocks presented within a session and the contents of each block are described in the *Method* sections of the specific experiments. Sessions differed from each other only in the order in which the blocks were presented. Within a session, no breaks were given, nor did participants receive any feedback on their performance. Thus, the completion of a block was immediately followed by the block information preceding the next block. Between sessions, breaks of about 5 min were given, during which the participant received feedback on mean RTs and mean PEs.

Data Analysis

Data were pooled across sessions. The first two trials of each block were discarded, as were RTs deriving from trials on which the participant did not respond correctly. Condition means of RTs and PEs were subjected to separate analyses of variance (ANOVAs). The multivariate solution is reported whenever it deviated from the univariate solution, with F values corresponding to Wilks's lambda (e.g., Stevens, 1992).

Experiment 1

Experiment 1 served to establish the phenomenon of asymmetric mixing costs by replicating and extending results reported by Van Duren and Sanders (1988). Van Duren and Sanders presented intact and noise-degraded digits for speeded identification in the pure/mixed design and observed mixing costs that were larger for intact digits than for degraded digits. Apart from attempting to replicate this result, I examined three issues. First, I examined whether the finding of asymmetric mixing costs generalizes to another context in which noise-degraded digits are replaced by segment-deleted digits. Second, I assessed the effects of the speed–accuracy trade-off, which Van Duren and Sanders failed to report. Third, I analyzed two types of sequential effects: nominal and category. Regarding the nominal sequence, a trial was classified as a repetition when the digit on that trial was the same as that on the preceding trial (regardless of the correspondence of the stimulus category) and as an alternation otherwise. Regarding the category sequence, a trial was classified as a repetition when the stimulus category on that trial was the same as that on the preceding trial (regardless of the correspondence of the digit) and as an alternation otherwise.

I examined category sequence to verify the exogenous nature of mixing costs, as discussed in the introduction. I examined nominal sequence for the following reason: It is a well-established finding that responding is faster to nominal repetitions than to nominal alternations (e.g., Bertelson, 1961; Campbell & Proctor, 1993; Kornblum, 1973; Pashler & Baylis, 1991). Pashler and Baylis (1991, Experiments 4 and 5) demonstrated that the size of this effect is larger when consecutive stimuli are not only nominally the same but also physically the same (i.e., when the stimuli are identical). For instance, they found that identical letters showed a larger advantage of repetition than letters that were nominally the same but differed in case or color. An illustration of how this differential effect might contribute to asymmetric mixing costs can be found in the design used by Van Duren and Sanders (1988). They used four alternative digits, with one instance for each intact digit and three instances for each degraded digit, constituted by different noise patterns. It follows that the probabilities of an identical repetition (i.e., the repetition of an instance) in the pure intact, pure degraded, mixed intact, and mixed degraded conditions were (1:4=) .25, [1:(4 × 3)=] .083, (.25/2=) .125, and (.083/2=) .0415, respectively. Therefore, the proportion of identical repetitions showed a greater drop from pure to mixed blocks for intact digits (from .25 to .125) than for degraded digits (from .083 to .0415), which might well underlie the greater mixing costs for intact than for degraded digits. This plausible mechanism would considerably reduce the theoretical significance of asymmetric mixing costs and therefore deserves closer examination.

Method

There were 24 participants, 12 in each of two groups. One group was presented with intact versus noise-degraded digits (Figure 1a vs. Figure 1b) in the pure/mixed design, whereas the other group was presented with intact versus segment-deleted digits (Figure 1a vs. Figure 1e). Each block contained 64 trials. In each of the four consecutive sessions, four blocks were presented, two pure blocks of either stimulus category and two randomly mixed blocks, both with a 1:1 ratio of occurrence of either stimulus category. The number of instances occurring within blocks was not equated across blocks. That is, 4, (4 × 4=) 16, and (16 + 4=) 20 instances occurred in the pure intact, pure degraded, and mixed blocks, respectively. In all other aspects, the method was the same as that described in the General Method section.

Results

Figure 2 shows the mean RTs and mean PEs as a function of the between-subjects factor of context (noise degraded or segment deleted) and the within-subjects factors of stimulus category (intact or degraded), block type (pure or mixed), and nominal sequence (repetition or alternation). The ANOVA on RTs indicated that there were no significant effects involving context. With respect to the within-subjects factors, the ANOVA yielded the following results: Mixing costs occurred in that responding was faster in pure blocks than in mixed blocks (526 vs. 536 ms), $F(1, 22) = 13.43$, $MSE = 398.03$, $p < .01$. Furthermore, responses were faster for intact digits than for degraded digits (494 vs. 568 ms), $F(1, 22) = 152.20$, $MSE = 1{,}720.70$, $p < .001$, and for nominal repetitions than for nominal alternations (511 vs. 551 ms), $F(1, 22) = 54.29$, $MSE = 1{,}430.24$, $p < .001$. Two 2-way interactions were found. First, an interaction between the effects of block type and nominal sequence, $F(1, 22) = 9.741$, $MSE = 173.11$, $p < .01$, indicated that mixing costs were greater for alternations (16 ms) than for repetitions (5 ms). Second, an interaction between the effects of category and block type, $F(1, 22) = 12.32$, $MSE = 325.07$, $p < .01$, indicated that mixing costs were more pronounced for intact digits (20 ms) than for degraded digits (1 ms). This interaction did not depend on nominal sequence ($F < 1$). When the ANOVA was conducted only on nominal alternations, mixing costs were still much larger for intact digits (25 ms) than for degraded digits (8 ms), $F(1, 22) = 11.33$, $MSE = 139.08$, $p < .01$.

The ANOVA performed on PEs showed that more errors occurred in the noise-degraded group than in the segment-deleted group (4.28% vs. 2.89%), $F(1, 22) = 7.16$, $MSE = 12.90$, $p < .05$. Furthermore, more errors occurred for degraded digits than for intact digits (4.61% vs. 2.56%), $F(1, 22) = 23.35$, $MSE = 8.65$, $p < .001$, as well as for alternations than for repetitions (3.98% vs. 3.18%), $F(1, 22) = 6.84$, $MSE = 4.48$, $p < .05$. More errors occurred in mixed blocks than in pure blocks (3.83% vs. 3.33%), but

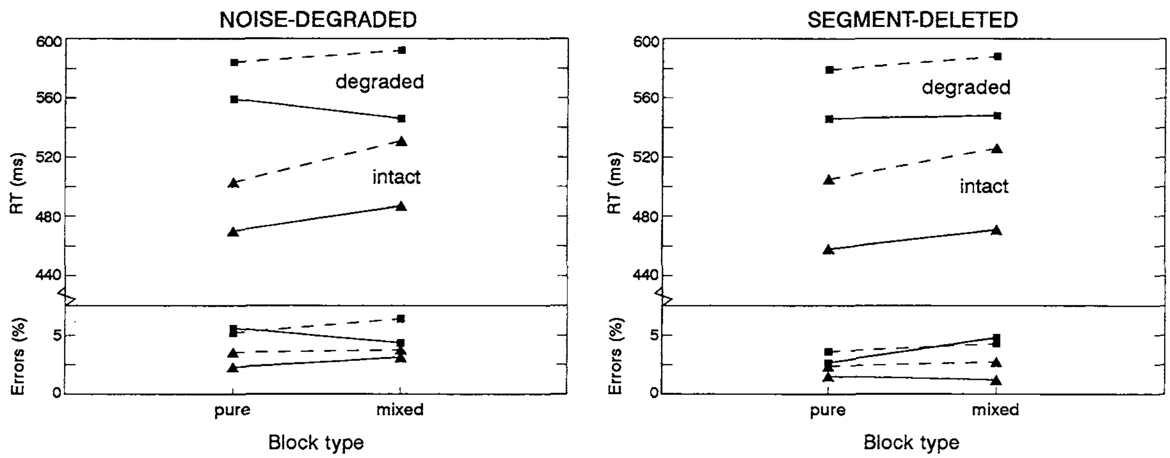

Figure 2. Mean reaction times (RTs) and mean error proportions in Experiment 1 as a function of stimulus category, block type, and nominal sequence, presented separately for the noise-degraded and the segment-deleted groups. Solid lines represent nominal intertrial repetitions; dashed lines represent nominal intertrial alternations.

these mixing costs failed to reach significance, $F(1, 22) = 2.09$, $MSE = 5.64$, $p = .16$. Finally, there was a significant four-way interaction involving all factors, $F(1, 22) = 5.47$, $MSE = 3.78$, $p < .05$. To examine this complex interaction, I performed separate ANOVAs on repetitions and alternations, with block type and stimulus category as within-subject factors and context as a between-subjects factor. The ANOVA on alternations did not yield any significant interaction. The ANOVA on repetitions showed a significant three-way interaction involving all factors, $F(1, 22) = 6.48$, $MSE = 6.48$ $p < .05$. This interaction indicated that for repetitions in the noise-degraded group, mixing costs were stronger for intact digits (0.87%) than for degraded digits (−1.23%), whereas this was reversed for repetitions in the segment-deleted group (−0.30% and 2.15% for intact and degraded digits, respectively).

Finally, category sequential effects were analyzed for the intact category in mixed blocks. For this purpose, each trial n in mixed blocks on which an intact digit occurred was categorized as follows: (a) second-order repetition, if the digits on trials $n - 1$ and $n - 2$ were both intact; (b) first-order repetition, if the digit on trial $n - 1$ was intact and that on trial $n - 2$ was degraded; (c) first-order alternation, if the digit on trial $n - 1$ was degraded and that on trial $n - 2$ was intact; and (d) second-order alternation, if the digits on trials $n - 1$ and $n - 2$ were both degraded. Figure 3 shows mean RTs and mean PEs as a function of category sequence. The negative and positive signs of the labels on the x-axis denote alternation and repetition, respectively, whereas the digit denotes the sequential order. An ANOVA on RTs, with category sequence as a within-subjects factor and degradation type as a between-subjects factor, revealed that the effect of category sequence was significant, $F(3, 20) = 4.04$, $p < .05$. Polynomial analysis of this effect revealed a strong linear component, $F(1, 22) = 12.4$, $MSE = 3{,}364.09$, $p < .01$, indicating that RTs gradually decreased from second-order alternation to second-order repetition. However, a paired t test showed that RTs for intact stimuli in pure blocks were still significantly shorter than RTs for second-order repetitions in mixed blocks, $t(23) = 2.95$, $p < .01$. An ANOVA on PEs, with the same factor structure as that on RTs, did not show any significant effect.

Discussion

Experiment 1 replicated the finding reported by Van Duren and Sanders (1988) of substantial mixing costs for intact stimuli when presented with degraded stimuli in the pure/mixed design. Three additional findings showed that these mixing costs were robust. First, these mixing costs were insensitive to the context constituted by the type of degradation. Second, they were not explicable in terms of a

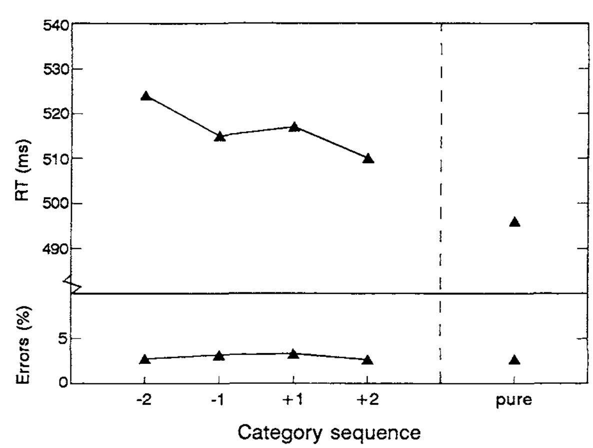

Figure 3. Mean reaction times (RTs) and mean error proportions for intact digits in Experiment 1 as a function of the intertrial sequence of stimulus category. On the x-axis, −2 and −1 denote second- and first-order alternations, respectively, and +1 and +2 denote first- and second-order repetitions, respectively.

speed–accuracy trade-off. Third, they were not attributable to a greater incidence of identical repetitions in pure blocks than in mixed blocks. It even appeared that mixing costs were larger for nominal alternations than for nominal repetitions. So, rather than causing mixing costs, nominal repetitions seem to compensate for the detrimental effects of mixing. In further agreement with the results of Van Duren and Sanders, the results of Experiment 1 show that mixing costs for degraded stimuli were much smaller than those for intact stimuli. This asymmetry of mixing costs also proved robust in view of its independence of degradation type and nominal sequence. The status of the asymmetry could be questioned, though, on the grounds of possible differential effects of a speed–accuracy strategy for intact and degraded stimuli. In particular, in the group receiving the segment-deleted digits, mixing costs on PEs proved more pronounced for degraded digits than for intact digits, which might balance the opposite pattern of mixing costs observed on RTs. However, this effect was present only for nominal repetitions, not for nominal alternations. This latter finding is important because mixing costs for nominal alternations are the least liable to possible explanations beyond the scope of the present article. For instance, they disconfirm the hypothesis suggested in the introduction to Experiment 1 that the asymmetry of mixing costs reflects a stronger reduction of the proportion of identical repetitions for intact digits than for degraded digits when going from pure to mixed blocks. From the perspective of the present study, then, Experiment 1 places the phenomenon of asymmetric mixing costs on a solid empirical footing.

Furthermore, Experiment 1 provides evidence that the mixing costs for intact digits are at least partially exogenous in nature by showing an advantage of category repetitions over category alternations. This effect was small and could not account for all the mixing costs because responding to intact digits in pure blocks was still faster than responding to intact digits preceded by two intact digits in mixed blocks. This could mean that the claim of an exclusively exogenous nature of mixing costs (Los, 1997; Maljkovic & Nakayama, 1994) is not justified. It could also mean, though, that performance in mixed blocks settles only to the level of that in pure blocks after an even longer sequence of category repetitions. In fact, results obtained by Maljkovic and Nakayama (1994) indicate that this may be the case. In their Experiment 7, the unique color of a target stimulus in a multiple stimulus display was either fixed (pure blocks) or variable (mixed blocks) across trials. They found that performance in mixed blocks gradually reached that in pure blocks across a sequence of about eight trials, during which the color of the target remained constant. Similarly, it is also possible that in Experiment 1, a further reduction in RTs would have shown up for sequences of intact digits beyond the second order, but a low reliability of the ensuing data discouraged such an analysis. Still, together with Los's (1997) observation that participants were incapable of reducing RTs in mixed blocks when they were informed on each trial of the forthcoming category, the present data suggest a major contribution of exogenous control.

Experiment 2

Having shown the reality of asymmetric mixing costs for stimuli of different perceptual categories, I now explore the underlying mechanism of this phenomenon. First, recall how the data of Experiment 1 are accounted for in terms of deadline tuning. In pure blocks, the position of the deadline is tuned in accordance with the time demands of either intact or degraded digits. On the basis of the data of Experiment 1, the deadline for degraded digits was used in mixed blocks because pronounced mixing costs were observed for intact digits but hardly any for degraded digits. This position of the deadline makes sense, because a sharper deadline may have caused an unacceptable error rate for the slow category, that is for degraded digits. Incidentally, the less pronounced mixing costs for nominal repetitions than for nominal alternations also fit this explanation well. The idea is that in pure blocks, the deadline is suboptimally placed for repetitions, because its position is predominantly determined by the slower alternations. Hence, a further relaxation of the deadline attributable to adding digits of a different category may not be so detrimental for repetitions as for alternations.

Next, recall how an alternate processing account would deal with the results of Experiment 1. Such an account assumes that at least partially different processing takes place on degraded digits and intact digits. Thus, Van Duren and Sanders (1988) suggested that in addition to processes involved in identifying intact digits, noise-degraded digits need a "clean-up" process that is capable of separating relevant from irrelevant features. Similarly, for segment-deleted digits a "fill-in" process can be proposed that fills in lacking segments. I would like to emphasize, though, that these labels merely serve explanatory purposes and that it is not essential that anything like cleaning up or filling in really take place on visual percepts. Rather, the central assumption of the alternate processing view is that shifting from one process to another takes time.[2] Thus, the mixing costs of Experiment 1 are attributed to frequent shifting in mixed blocks among "fast" processes for intact digits and additional clean-up or fill-in processes for degraded digits. According to the mental inertia hypothesis (Allport et al., 1994; Los, 1996), the asymmetry of these mixing costs indicates that clean-up or fill-in processes demand more control than the fast processes for intact digits, which makes shifting from degraded to intact digits harder than vice versa. Furthermore, the smaller mixing costs for nominal repetitions than for nominal alternations may indicate that (partially) reviewing the same digit compensates for the detrimental effects of shifting.

The failure to find effects of a speed–accuracy trade-off in Experiment 1 might be taken as evidence against a criterion

[2] As the higher order category sequential effects of Experiment 1 indicate, a two-state model, implying that the perceptual system is either optimally prepared for processing intact digits or for processing degraded digits, is probably too simple. A more realistic view is that the presentation of a degraded digit may interfere proactively across several trials with the processing of later intact digits (cf. Allport et al., 1994).

account of mixing costs, but, as argued before, this may have been due to the small size of mixing costs or the unreliability of the accuracy measure. Experiment 2, therefore, provided a test for the proposed mechanisms whereby evidence is reflected in RTs. For this purpose, the noise-degraded and segment-deleted digits were presented in the pure/mixed design. Because the data of Experiment 1 showed almost equal RTs for these stimulus categories, a single deadline should satisfy either category. Hence, according to a deadline account, RTs for noise-degraded and segment-deleted digits should be insensitive to manipulations of block type because these categories are subject to a single criterion across blocks. By contrast, an alternate processing account still predicts the occurrence of mixing costs because noise-degraded and segment-deleted categories are assumed to call on different processing. In addition, in Experiment 2 I examined more generally than in Experiment 1 the influence of different numbers of instances occurring in pure and mixed blocks. One group of participants (the "unbalanced" group) received in mixed blocks all noise-degraded and segment-deleted instances, yielding twice as many instances in mixed blocks than in pure blocks, whereas another group (the "balanced" group) received as many instances in mixed blocks as in pure blocks. If the number of instances somehow contributes to mixing costs, the unbalanced group should show greater mixing costs than the balanced group.

Method

There were 40 participants, 20 in each of two groups. One participant exceeded a fixed maximal PE of 15% in several conditions and was therefore replaced by another participant. All participants received noise-degraded and segment-deleted digits (i.e., Figure 1b vs. Figure 1e) presented in the pure/mixed design. Each block contained 64 trials. In each of the four sessions, four blocks were presented: two pure blocks of either category and two mixed blocks, both with a 1:1 ratio of occurrence of either category. The unbalanced group received (4 × 4=) 16 instances in each pure block and (16 + 16=) 32 instances in each mixed block, whereas the balanced group received 16 instances in all blocks using the balancing procedure described in the General Method section under *Block Type*. All other details were as described in the General Method section.

Results

Figure 4 shows mean RTs and mean PEs as a function of the between-subjects factor of instance balance (balanced or unbalanced) and the within-subject factors of stimulus category (noise degraded or segment deleted) and block type (pure or mixed). An ANOVA on these RT data indicated that the observed 26-ms difference between the groups of instance balance was not significant, $F(1, 38) = 1.55$, $MSE = 17{,}969.173$, $p > .20$, because it was mainly attributable to the relatively slow responding of 4 participants in the unbalanced group. The effect of block type was highly significant, $F(1, 38) = 61.85$, $MSE = 254.20$, $p < .001$, reflecting the fact that RTs in pure blocks (594 ms) were shorter than RTs in mixed blocks (614 ms). These mixing costs were not due to a larger number of instances in mixed blocks than in pure blocks because the interaction between the effects of instance balance and block type was insignificant, $F(1, 38) = 1.02$, $MSE = 254.20$, $p > .30$. Furthermore, RTs for noise-degraded digits (601 ms) were somewhat shorter than RTs for segment-deleted digits (608 ms), but not significantly so, $F(1, 38) = 1.87$, $MSE = 1{,}096.40$, $p > .15$. Contrary to the prediction of a deadline account, this result suggests that a difference in RTs between categories is not a necessary condition for the occurrence of mixing costs. Regarding PEs, the ANOVA yielded only a significant main effect of stimulus category, $F(1, 38) = 14.33$, $MSE = 5.09$, $p < .01$, indicating that errors were more frequent for noise-degraded digits (4.41%) than for segment-deleted digits (3.06%).

In pure blocks, remarkable individual differences were observed on the relative RTs for the two stimulus categories. Whereas some participants responded considerably faster to segment-deleted digits than to noise-degraded digits, this was reversed for other participants. This suggests the

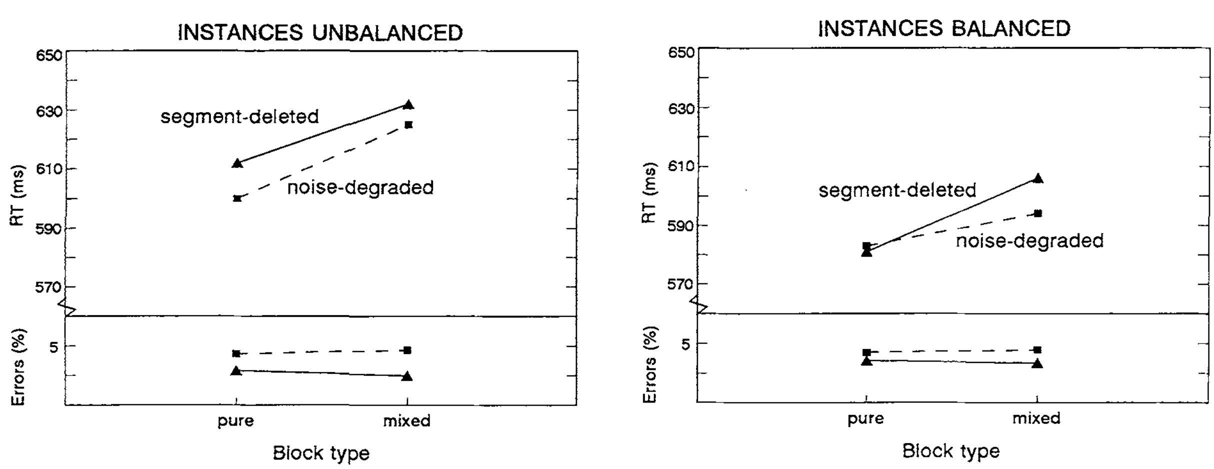

Figure 4. Mean reaction times (RTs) and mean error proportions in Experiment 2 as a function of stimulus category, block type, and instance balance.

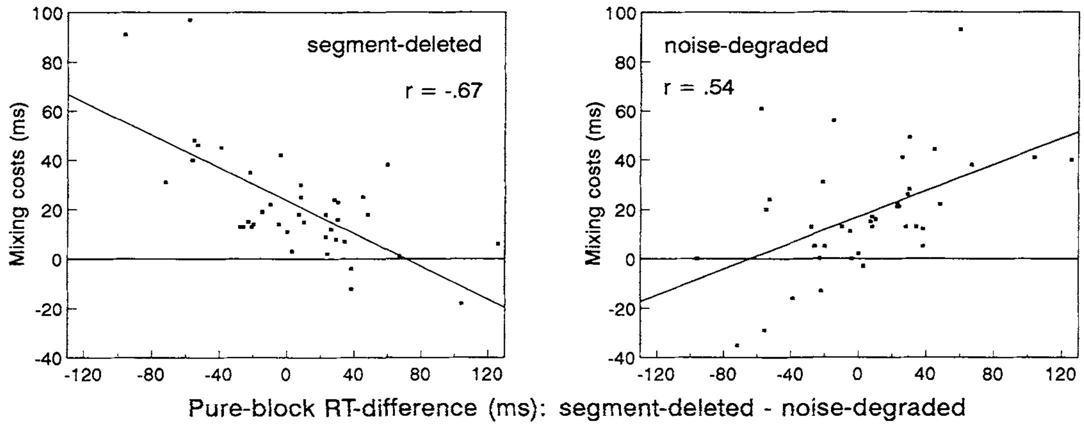

Figure 5. Individual mixing costs for the noise-degraded and segment-deleted category related to individual pure-block reaction time (RT) difference in Experiment 2. Pure-block RT difference is defined as the time by which noise-degraded stimuli are responded to faster than segment-deleted stimuli in pure blocks.

possibility that averaging across these individual differences concealed possible interactions between the effects of block type and stimulus category in the ANOVA reported earlier. Figure 5 therefore plots, for each participant and both stimulus categories, mixing costs as a function of the *pure-block RT difference,* defined as the time by which noise-degraded digits are responded to faster than segment-deleted digits in pure blocks. Thus, the data points with negative values of pure-block RT difference stem from participants who responded relatively quickly to segment-deleted digits, whereas the data points with positive values of pure-block RT difference stem from participants who responded relatively quickly to noise-degraded digits. From Figure 5 it is immediately apparent that mixing costs for a stimulus category increase as that category acquires the character of a fast category, whereas mixing costs become negligible (or even turn into benefits) as a category acquires the character of a slow category. This observation is statistically supported by a significant negative correlation between pure-block RT difference and mixing costs for segment-deleted digits ($r = -.67$), $t(38) = -5.57$, $p < .001$, and by a significant positive correlation between pure-block RT difference and mixing costs for noise-degraded digits ($r = .54$), $t(38) = 3.91$, $p < .001$. Furthermore, the correlation between pure-block RT difference and error mixing costs was significant for segment-deleted digits ($r = -.36$), $t(38) = -2.34$, $p < .05$, but not for noise-degraded digits ($r = .22$), $t(38) = 1.36$, $p > .15$. Although the latter correlations were not strong, they run counter to the predictions of the speed–accuracy trade-off. If anything, error mixing costs tended to increase as a category acquired the character of a fast category.

Finally, category sequential effects were analyzed to verify whether the observed mixing costs were exogenous in nature. Analogous to the procedure followed in Experiment 1, first- and second-order alternations, and first- and second-order repetitions, were derived for both the noise-degraded and the segment-deleted categories. Figure 6 shows mean RTs and mean PEs as a function of category sequence and stimulus category. An ANOVA on these RTs (with instance balance as a between-subjects factor) showed main effects of stimulus category, $F(1, 38) = 4.81$, $MSE = 1{,}425.87$, $p < .05$, and category sequence, $F(3, 36) = 8.21$, $p < .001$. Polynomial analysis of the latter effect showed a strong linear component, $F(1, 38) = 23.74$, $MSE = 420.41$, $p < .001$, indicating progressively faster responding from second-order alternations to second-order repetitions. Paired t tests indicated that responding in pure blocks was still faster than responding for second-order repetition trials in mixed blocks, both for segment-deleted stimuli, $t(39) = 4.54$, $p < .001$, and for noise-degraded stimuli, $t(39) = 2.48$, $p < .05$. An ANOVA on PEs, with the same factor structure as that on RTs, yielded only a significant effect of stimulus category, $F(1, 38) = 22.54, MSE = 9.78, p < .001$.

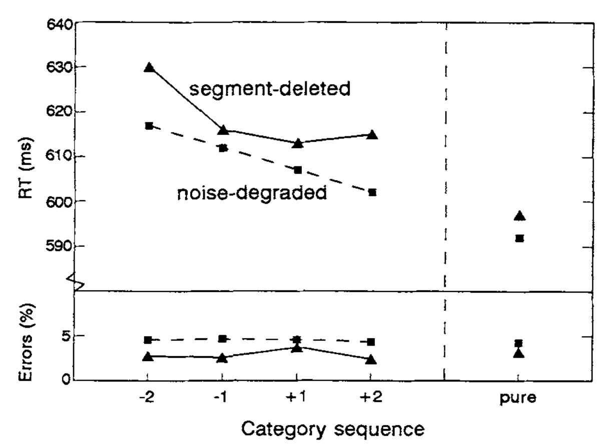

Figure 6. Mean reaction times (RT) and mean error proportions in Experiment 2 as a function of stimulus category and the intertrial sequence of stimulus category. On the x-axis, -2 and -1 denote second- and first-order alternations, respectively, and $+1$ and $+2$ denote first- and second-order repetitions, respectively.

Discussion

The results of Experiment 2 show that the number of instances occurring in a block is probably not an important determinant of mixing costs. In the unbalanced group, twice as many instances occurred in mixed blocks as in pure blocks, but the mixing costs were hardly any larger than in the balanced group, where the number of instances was equated across blocks. This confirms the conclusion of Experiment 1 that mixing costs constitute a robust phenomenon that is not an artifact of a higher probability of instance repetitions in pure blocks than in mixed blocks.

More important was the demonstration that clear mixing costs occurred even though noise-degraded and segment-deleted digits showed about equal average RTs in pure blocks. I anticipated that this finding would disconfirm a deadline account of mixing costs, reasoning that participants would use the same deadline in mixed blocks as in pure blocks. This logic was weakened, however, by the unexpected individual differences in the relative pure-block RTs for noise-degraded and segment-deleted digits, due to which many participants featured a fast category and a slow one. Taking this into account, the pattern of mixing costs in Experiment 2 proved highly similar to that in Experiment 1: The fast stimulus category showed clear mixing costs, but the slow category showed hardly any. As argued before, this pattern of data fits a deadline account quite well.

On closer examination, at least two observations argue against a deadline account, however. First, a criterion account predicts that mixing costs do not occur when two stimulus categories need about equal processing times in pure blocks. However, Figure 4 shows that the pure-block RT differences close to 0 ms, say from −20 ms to 20 ms, are almost without exception associated with mixing costs regardless of stimulus category. This is also expressed by the fact that the best-fitting regression lines in Figure 4 intersect the x-axis remotely from the origin. Second, there was a clear trend for mixing costs on RTs to be positively correlated with mixing costs on accuracy. At the very least, this means that the extra processing time used by the fast category in mixed blocks relative to pure blocks could not be used to acquire a higher precision, as predicted by a criterion account.

Regarding category sequential effects, these results replicate those of Experiment 1, showing progressively faster responding with a decreasing order of category alternation and an increasing order of category repetition. As was argued in the *Discussion* section in Experiment 1, this effect is consistent with the view that the observed mixing costs have an exogenous source.

Experiment 3

The results of Experiment 2 rendered some, but not yet convincing, support for an alternate processing account of the mixing costs under examination. My goal in Experiment 3 was to obtain more direct evidence. This experiment used 3:1 and 1:3 proportions of occurrence of noise-degraded and segment-deleted digits in mixed blocks, in addition to the 1:1 proportion used in Experiment 2. This resulted in four conditions of category frequency: (a) 100% for either category in pure blocks, (b) 75% for the frequent category in the mixed-3:1 and mixed-1:3 blocks, (c) 50% for either category in the mixed-1:1 blocks, and (d) 25% for the infrequent category in the mixed-1:3 and mixed-3:1 blocks. In effect, category frequency manipulated the intertrial variability of a block.[3] That is, as a category is presented less frequently in a block, the trials on which that category occurs are more often preceded by trials on which the other category occurs. According to an alternate processing view, the present mixing costs are attributable to a time loss in switching between processes such as cleaning up and filling in. Therefore, this account predicts that mixing costs increase in accordance with the proportion of alternation trials. By contrast, accounts that deny the involvement of differential processing do not so readily predict that RT is affected by the proportion of alternation trials. Although Experiment 2 showed that deadline tuning may account for mixing costs when there is a difference in pure-block RT, there is no reason to predict a further relaxation of the deadline for a mixed-3:1 block as compared with a mixed-1:1 block.

Method

There were 12 participants, 1 of whom failed to respond within the fixed maximum of 1,500 ms on at least 10% of the trials of several conditions and was therefore replaced by another participant. As in Experiment 2, noise-degraded and segment-deleted digits (Figure 1b vs. Figure 1e) were presented in the pure/mixed design. Block type had four levels: pure, mixed-3:1, mixed-1:1, and mixed-1:3. In all blocks, 16 instances occurred with an equal probability of being presented on each trial (cf. the instance balancing procedure described in the General Method section). Participants were tested in four consecutive sessions, each comprising five blocks: two pure blocks of either stimulus category and three mixed blocks with 3:1, 1:1, and 1:3 proportions of occurrence of either category. Each block contained 48 trials. In all other respects, the method adhered to the details described in the General Method section.

Results

Figure 7 shows the mean RTs and mean PEs as a function of stimulus category (noise degraded or segment deleted) and category frequency (100%, 75%, 50%, or 25%). An ANOVA on the RT data revealed a significant main effect of category frequency, $F(3, 9) = 10.30$, $MSE = 1{,}502.62$, $p < .01$. This effect reflected a gradual increase in RTs of a category as that category occurred less frequently in a block. Relative to the 50% condition, 8 of the 12 participants showed costs in RTs in the 25% condition for both catego-

[3] A priori, category frequency may also be conceived of as manipulating the probability that participants endogenously prepare for one or the other category. However, in a previous study (Los, 1997) I found that informing the participant at the start of a trial on the forthcoming category did not result in any reduction of mixing costs. Therefore, this endogenous preparation view is not a promising candidate to account for frequency manipulations in mixed blocks.

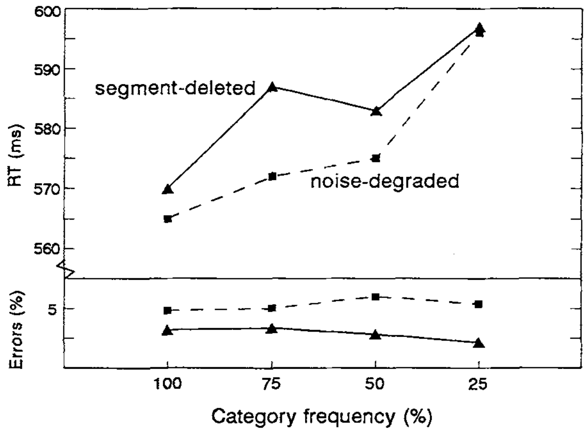

Figure 7. Mean reaction times (RTs) and error proportions in Experiment 3 as a function of stimulus category and category frequency.

ries. One participant showed costs for one category and neither costs nor benefits for the other category. The remaining 3 participants showed costs for one category and benefits of 3, 5, and 43 ms for the other category. Therefore, of 24 observations (12 participants × 2 categories), there was only one clear violation of the general trend that responding was slower in the 25% condition than in the 50% condition. Regarding PEs, the only significant effect was a main effect of stimulus category, $F(1, 11) = 6.80$, $MSE = 20.69$, $p < .05$, indicating that responding was less accurate to noise-degraded than to segment-deleted digits.

Discussion

The results of Experiment 3 show that RTs of a stimulus category increase as its incidence decreases. Most notably, responding was clearly slower in the 25% condition than in the other frequency conditions. This result is predicted by an alternate processing account because stimuli of the infrequent category are relatively often preceded by stimuli of the other category. This requires a shift in processing, which in turn results in a loss of time. It is far less obvious why a deadline account would predict the slow responding in the 25% condition. In the case of a pronounced pure-block RT difference between the stimulus categories, this result could be predicted, though, for the fast category. In that case, the slow category would contribute more strongly to the tuning of the deadline as it is more frequently presented, thus making the deadline for the fast category increasingly less appropriate. However, the observed slow responding in the 25% condition proved far too general to justify this account. This result held just as well for the slow category as for the fast category and irrespective of whether participants showed a clear pure-block RT difference. It is because of this generality that deadline tuning provides a poor account of the present results.

Experiment 4

The pairs of stimulus categories used in Experiments 1–3 proved effective in producing mixing costs. It has been assumed thus far that this was due to different computational processing called on by the different stimulus categories. To test this account, one must demonstrate that mixing costs do not occur when different stimulus categories do not require computationally different processing. Therefore, in Experiment 4 I used stimulus categories that differed on features that seem trivial from a computational point of view. The categories were all variations within the noise-degraded category. One group of participants, the noise group, received noise-degraded categories with either 8 or 12 noise dots. A pilot study showed that participants responded more slowly to the category with 12 noise dots than to that with 8 noise dots. Therefore, the predictions for this group were precisely the opposite of those in Experiment 3. A deadline account would predict mixing costs for the category with 8 noise dots because responding to this fast category takes place at a less conservative criterion in pure blocks than in mixed blocks. By contrast, an alternative processing account predicts the absence of mixing costs in this group because both categories are assumed to call on the same computational processes. Another group of participants, the form group, received noise-degraded digits either within a rectangular frame or within a round frame. This tested whether mixing costs would occur for any salient perceptual manipulation. In that case, mixing costs would merely result from a time-consuming orientation response (e.g., Sokolov, 1963) rather than from switching among computational processes.

Method

There were 24 participants, 12 in each of two groups. Two participants (1 in each group) failed to respond on at least 10% of the trials in several conditions and were therefore replaced by 2 other participants. All participants received the noise-degraded digits with 12 noise dots within a rectangular frame (i.e., Figure 1b) as one of the stimulus categories presented in the pure/mixed design. The other category was the noise-degraded category with 8 noise dots (i.e., Figure 1c) for the noise group and the noise-degraded category surrounded by a round frame (i.e., Figure 1d) for the form group. In all blocks, 16 instances occurred, using the instance balancing procedure described in the General Method section. In the form group, the instances of both categories had virtually identical noise patterns (cf. the General Method section). This enabled the construction of mixed blocks in which the excluded instances of one category were replaced by the corresponding instances of the other category. The method was identical to the method of Experiment 3 in all other respects.

Results

Figure 8 shows RTs and PEs as a function of the between-subjects variable of group (noise or form) and the within-subjects variables of stimulus category (8 or 12 noise dots; rectangular or circular frame) and category frequency (100%, 75%, 50%, or 25%). Regarding RTs, a general

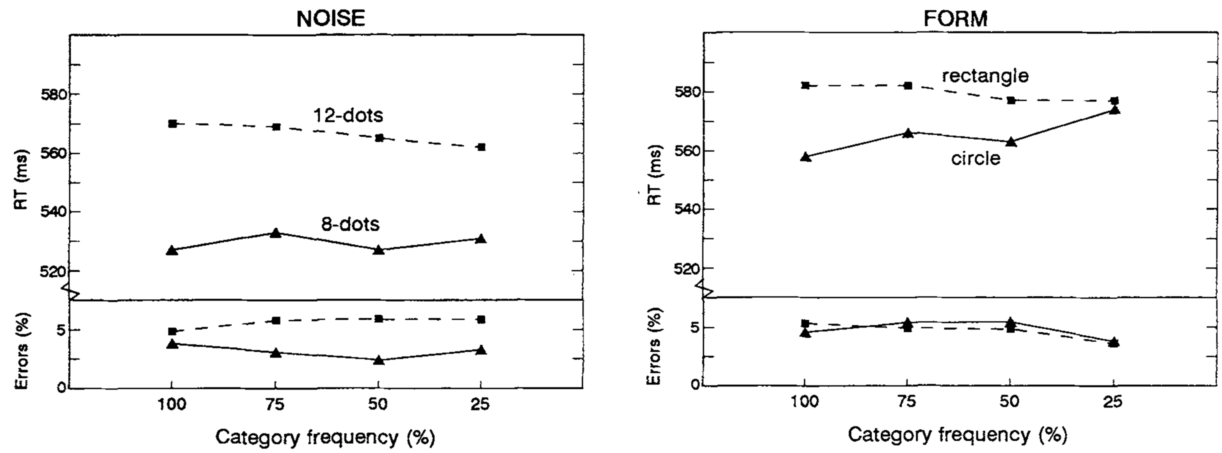

Figure 8. Mean reaction times (RTs) and error proportions in Experiment 4 as a function of stimulus category and category frequency, presented separately for the noise group and the form group.

ANOVA involving these three factors revealed a main effect of stimulus category, $F(1, 22) = 60.41$, $MSE = 521.35$, $p < .001$, indicating slower responding to noise-degraded digits within a rectangular frame and 12 noise dots than to digits belonging to the other categories. This effect was stronger in the noise group than in the form group, $F(1, 22) = 11.67$, $MSE = 521.35$, $p < .01$, although the main effect of category proved significant in both groups when analyzed in separate ANOVAs: $F(1, 11) = 45.46$, $MSE = 717.92$, $p < .001$ in the noise group; $F(1, 11) = 15.23$, $MSE = 324.78$, $p < .01$ in the form group. The significance of the effect of stimulus category in the form group was surprising because it indicates that an apparently trivial feature such as the form of the surrounding frame may still affect RTs. The most important result, however, is that neither the general ANOVA nor the separate ANOVAs on the two groups revealed any tendency toward a main effect of category frequency: $F(3, 20) = 1.16$, $p > .30$ in the general ANOVA; Fs < 1 in each group when tested separately. Finally, the general ANOVA revealed a significant interaction between the effects of stimulus category and category frequency, $F(3, 20) = 3.27$, $p < .05$, reflecting that RTs for the slow category decreased as the category occurred less frequently, whereas RTs for the fast category increased. This effect was modest and failed to reach statistical significance in the separate ANOVAs: $F(3, 9) = 2.70$, $p > .10$ in the form group; $F(3, 9) < 1$ in the noise group.

Regarding PEs, the general ANOVA yielded a main effect of stimulus category, $F(1, 22) = 8.71$, $MSE = 7.45$, $p < .01$, which interacted with the effect of group, $F(1, 22) = 10.76$, $MSE = 7.45$, $p < .01$. Separate ANOVAs on the two groups specified this interaction, indicating that responding was less accurate to the 12-dot degraded category within a rectangular frame than to the 8-dot degraded category, $F(1, 11) = 17.27$, $MSE = 8.38$, $p < .01$, but not less accurate than responding to the degraded category within a round frame ($F < 1$).

Discussion

The data from Experiment 4 contrast sharply with those from Experiment 3. In Experiment 3, mean RTs for both stimulus categories strongly depended on the incidence of that category in a block; in Experiment 4, this dependence was gone. Together, these data demonstrate that mixing costs cannot be elicited by just any perceptual manipulation. Rather, they suggest that different computational processing demands of the stimuli constitute a necessary condition for the occurrence of mixing costs. In particular, it cannot be maintained that the mixing costs observed in Experiment 3 were attributable to a pure-block RT difference between noise-degraded and segment-deleted digits because in that case this effect should have been manifest in the noise group of Experiment 4 as well. However, in spite of a pure-block RT difference of 38 ms between digits with 8 and 12 noise dots, mixing these categories in whatever proportion did not affect RTs. This result convincingly adds to the evidence against criterion accounts. If a criterion is determined by the slow or frequent category occurring in a block, digits with 8 noise dots should have shown a noticeable increase in RTs when combined with digits with 12 noise dots, especially when occurring on only 25% of the trials. The data did not show any tendency in this direction, however. Again, if salient variations of form had caused the mixing costs in Experiment 3, this effect should have been present in the form group as well. However, although the shape of the frame unexpectedly affected RTs in pure blocks and as such cannot be considered as trivial, it did not cause general mixing costs. Instead, in mixed blocks the RTs for digits within round and rectangular frames tended to converge toward a common average. This observation suggests that slight variations in the processing criterion might have played a role in the form group of Experiment 4. In Experiment 3, by contrast, RTs merely increased as a category was presented less frequently regardless of stimulus category and pure-block RT difference. This general increase is consistent with the prediction of an alternate processing account.

Experiment 5

So far, some evidence has been obtained that mixing costs are attributable to shifting among different computational processes, but no effort has been invested to specify these processes in any detail. My aim in Experiment 5 was to take a first step in this direction. The computational processes encountered have been provisionally described in terms of cleaning up and filling in. Although not meant as a veritable specification of the contents of these processes, these descriptions do indicate at which level the processes are supposed to operate: not at an early level, where processing is bound to retinotopic coordinates, but at a later level, where processing takes place on some base representation (Marr, 1982; Ullman, 1984). As such, the computational processes may be identified with visual routines, which have been proposed to conduct a goal-directed analysis on the base representation to enable object recognition (Trick & Pylyshyn, 1994; Ullman, 1984). An alternative to this view localizes the origin of mixing costs at an earlier stage of visual processing, inspired by the observation that the stimulus pairs that produced mixing costs in Experiments 1–4 mutually differed in the local density of dots within the stimulus area. Thus, compared with intact digits, noise-degraded digits have relatively many dots within their frames, whereas segment-deleted digits have relatively few. It could be, then, that at an early stage of processing the visual system is sensitive to these density variations, such that a change in density requires adjustment.[4]

To contrast the competing hypotheses, in Experiment 5 I introduced a new stimulus category: dice (see Figure 9d). Dice were presented pairwise in the pure/mixed design with intact, noise-degraded, and segment-deleted digits. Dice have relatively few dots within their frames and are, in that respect, more similar to segment-deleted digits than to intact digits and more similar to intact digits than to noise-degraded digits. Therefore, the local density hypothesis predicts that dice will show increasing mixing costs when they are mixed with segment-deleted, intact, and noise-degraded digits, respectively. On the other hand, dice are intact, and in that respect are more similar to intact digits than to degraded digits. Therefore, the visual routines hypothesis predicts smaller mixing costs for dice when presented with intact digits than when presented with either segment-deleted or noise-degraded digits.

To reduce the likelihood that categories would be mutually confused, I presented digits within round frames and dice dots in square frames. In particular, this measure was meant to provide a useful cue to distinguish dice from segment-deleted digits, which may be somewhat similar otherwise. To the degree that this measure is insufficient, the mixing costs deriving from the combination of dice and noise-degraded digits enable one to disambiguate some anticipated conflicting results. Specifically, if category confusion is an important cause of mixing costs when dice are mixed with segment-deleted digits, then much smaller mixing costs are predicted when dice are mixed with the highly dissimilar noise-degraded digits.

To summarize, in Experiment 5 I compared mixing costs for dice when presented with segment-deleted digits, intact digits, and noise-degraded digits. The local density hypothesis predicts an increase in mixing costs across these three conditions, the confusion hypothesis a decrease, and the visual routines hypothesis a U-shaped function.

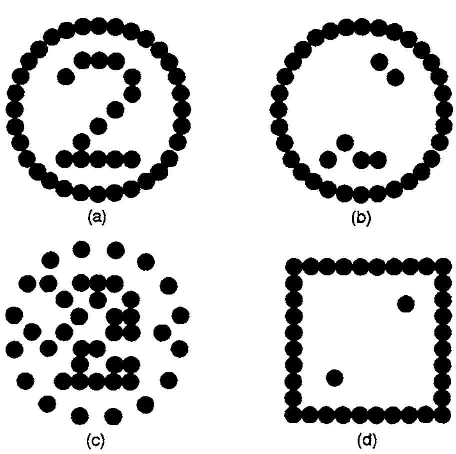

Figure 9. Examples of the stimulus categories, depicted as negatives, presented to the segment-deleted group (a, b, and d) and the noise-degraded group (a, c, and d) in Experiment 5. The stimuli depict the numerical value 2 represented by an intact digit (a), a segment-deleted digit (b), a noise-degraded digit (c), and dice (d).

Method

There were 30 participants, 15 in each of two groups. Both groups received three categories in the pure/mixed design: dice, intact digits, and degraded digits (see Figure 9). In the segment-deleted group, the degraded category comprised segment-deleted digits, whereas in the noise-degraded group, it comprised noise-degraded digits. All digits were presented within a round frame composed of 32 dots. This frame had a diameter of 5.7 cm and subtended a visual angle of 6.5° given a viewing distance of 50 cm. The dice dots were presented within a square, 5 × 5 cm frame of 36 dots that subtended a visual angle of 5.7°. Both digits and dice represented the numerical values 2, 3, 4, or 5, on which the participants based their response, as described in the introduction.

In mixed blocks, the presentation of the categories was pairwise, so that each category occurred in two of three different contexts: intact–degraded, intact–dice, and degraded–dice. To elicit sizable mixing costs, I used only mixed-3:1 and mixed-1:3 blocks in each context, in addition to the pure blocks. This yielded a total of nine

[4] This view could be related to the property of the visual system to analyze a percept into its constituent spatial frequencies at an early stage of processing (e.g., Blakemore & Campbell, 1969; Ginsburg, 1986). Thus, noise-degraded digits may require the visual system to discard the highest spatial frequencies, whereas segment-deleted digits may require the visual system to emphasize the high spatial frequencies. In this view, then, mixing costs are caused by shifts among spatial frequency channels. I am grateful to Andries Sanders for drawing my attention to this possibility.

different blocks: three pure, three mixed-3:1, and three mixed-1:3. In each of two consecutive sessions, 15 blocks were presented in random order, comprising the three pure blocks and each of the six mixed blocks twice. In the first session, the order of presentation of these blocks was counterbalanced across participants. In the second session, each participant received the blocks in the reverse order. The instance balancing procedure (cf. the General Method section) was applied to the degraded categories, but not to the intact and dice categories, which had only one instance for each of their numerical values. Each block contained 34 trials. All other details were as described in the General Method section.

Results

Figure 10 shows mean RTs and mean PEs as a function of stimulus category, category frequency, and context. The analysis of these data focused on three issues. The first issue is how a single category is affected by different contexts in the pure/mixed design. This issue is pertinent with respect to the hypotheses addressed earlier. The next issue is how different categories affect each other within a single context. This issue focuses on the asymmetry of mixing costs. A final issue is whether mixing costs within a single context can be predicted from mixing costs obtained in the other two contexts.

Context effects. To assess context effects, I performed six separate ANOVAs, one for each combination of category (3) and performance measures (i.e., RTs and PEs). In each ANOVA, group (segment deleted or noise degraded) served as a between-subjects factor and category frequency (3) and context (2) served as within-subjects factors. In all ANOVAs, identical data occurred in the pure-block cells (i.e., 100% category frequency) for the two different contexts in which a category occurred.

The three ANOVAs on RTs showed mixing costs in that a category was more slowly responded to when it occurred less frequently, $Fs(2, 27) = 46.67$, 36.63, and 39.84, $ps < .001$, for dice, intact digits, and degraded digits, respectively. No effect of context was found for the intact category, $F(1, 28) = 1.78$, $MSE = 147.11$, $p > .15$, indicating that RTs for intact digits were about equally affected by a context of dice and of degraded digits. By contrast, for dice and degraded digits, highly significant context effects were found, $Fs(1, 28) = 26.39$ and 40.56, $MSEs = 209.26$ and 410.87, $ps < .001$, respectively. These effects indicate that RTs for dice and degraded digits were longer when combined with each other than when combined with intact digits. In addition, both for dice and for degraded digits the effect of context interacted significantly with that of category frequency, $Fs(2, 27) = 15.63$ and 19.98, $ps < .001$, respectively. This indicates that for these categories, the differential effect of context increased as the category occurred less frequently. Finally, in none of the three ANOVAs was there a significant main effect of group, nor was there a significant interaction involving group. This indicates that the noise-degraded and segment-deleted categories behaved similarly, both in the way they were affected by different contexts and in the way they functioned as a context for other categories.

The three ANOVAs on PEs revealed a significant group effect for intact digits, $F(1, 28) = 5.03$, $MSE = 11.04$, $p < .05$, indicating that responding to these digits was slightly less erroneous in the noise-degraded group than in the segment-deleted group. Furthermore, there was a tendency toward fewer errors for intact digits in the context of degraded digits than in the context of dice, $F(1, 28) = 3.48$, $MSE = 3.51$, $p = .07$. Both for dice and for degraded digits a significant effect of frequency was found, $Fs(2, 27) = 3.77$

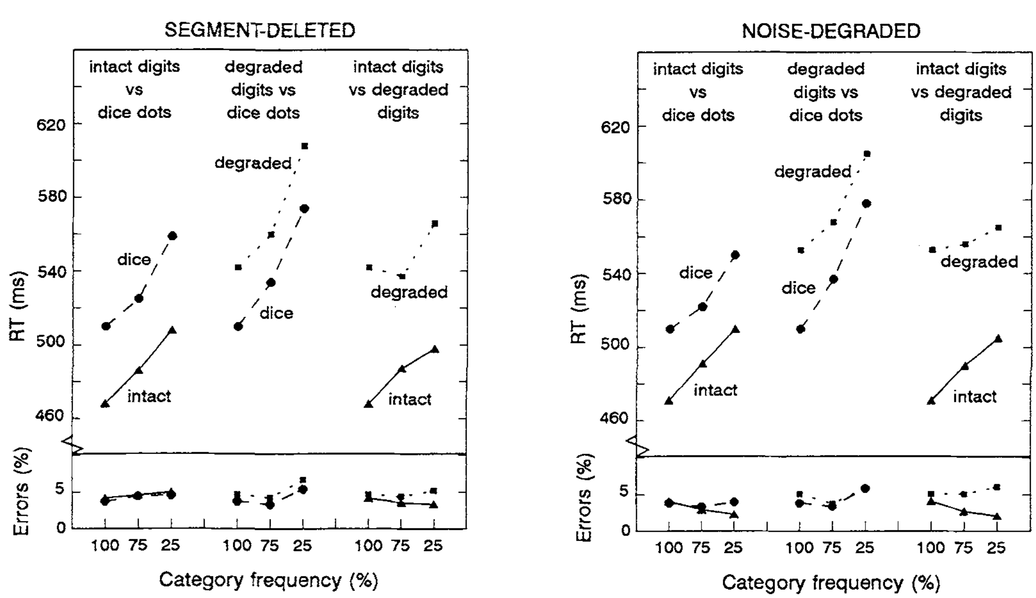

Figure 10. Mean reaction times (RTs) and error proportions in Experiment 5 as a function of stimulus category, context, and category frequency, separately for the segment-deleted and the noise-degraded groups.

and 5.20, *ps* < .05, respectively, indicating that these categories were less accurately responded to as they occurred less frequently.

Mixing-cost asymmetries. To assess mixing-cost asymmetries, I performed six separate ANOVAs, one for each combination of context (3) and performance measure (2). In each ANOVA, group served as a between-subjects factor and category and category frequency as within-subject factors.

For each context, the ANOVA on RTs showed a highly significant effect of category. Responding was faster to intact digits than to degraded digits, $F(1, 28) = 262.54$, $MSE = 758.50$, $p < .001$, to intact digits than to dice, $F(1, 28) = 92.84$, $MSE = 794.15$, $p < .001$, and to dice than to degraded digits, $F(1, 28) = 29.23$, $MSE = 1{,}556.64$, $p < .001$. Furthermore, responding was slower as categories occurred less frequently, $Fs(2, 27) = 77.02$, 77.18, and 33.10, $ps < .001$, in the intact/dice, dice/degraded, and intact/degraded contexts, respectively. An asymmetry of mixing costs occurred only in the intact/degraded context, as indicated by a significant interaction between the effects of category and category frequency, $F(2, 27) = 5.82$, $p < .01$.

The ANOVA on PEs in the dice/degraded context yielded a significant effect of frequency, $F(2, 27) = 8.19$, $p < .01$, mainly because of relatively inaccurate responding in the 25% frequency condition. An effect of group was found in the dice/intact context, $F(1, 28) = 5.01$, $MSE = 9.32$, $p < .05$, indicating that fewer errors occurred in the noise-degraded group than in the segment-deleted group. Because the intact/dice context was identical in both groups, this effect should probably be considered a Type I error. Finally, in the intact/degraded context, the ANOVA revealed a significant effect of category, $F(1, 28) = 15.23$, $MSE = 9.25$, $p < .01$, indicating that fewer errors occurred for intact digits than for degraded ones. Figure 10 suggests that this effect was modified by category frequency, but the interaction between category and category frequency fell short of significance, $F(2, 27) = 2.25$, $p > .10$.

Additivity of mixing costs. As is apparent both from Figure 10 and from the ANOVAs, mixing costs were clearly larger in the degraded/dice context than in either of the other contexts. This gives rise to the question of whether mixing costs within this context can be predicted from an additive combination of mixing costs from the other contexts. This would support a model that describes shifting in the degraded/dice context in two discrete stages: one stage for shifting to or from a routine such as filling in or cleaning up and another stage for shifting between routines involved in digit identification and dot enumeration. Under this model, the following equation holds:

$$\mathrm{RT(dice)}_{f,\mathrm{dice/degraded}} = \mathrm{RT(dice)}_{\mathrm{pure}} + \mathrm{MC(intact)}_{f,\mathrm{intact/degraded}} + \mathrm{MC(dice)}_{f,\mathrm{dice/intact}}$$

where MC denotes mixing costs, the first subscript f denotes frequency (75% or 25%), and the second subscript denotes the context. In this equation, the first term yields an estimate of RTs for dice under optimal circumstances, the second term yields an estimate of the costs for disengaging visual routines involved in specific processing on degraded digits (e.g., cleaning up or filling in), and the third term yields an estimate of costs for disengaging visual routines involved in general processing on digits (for the benefit of enumeration processes). The predictions of the model for degraded digits in the degraded/dice context are derived by substituting in the above equation "degraded" for "dice" and "dice" for "degraded." Figure 11 shows observed and predicted data in the degraded/dice context.

The discrete shift model was tested by contrasting the predicted data in the 75% and 25% frequency conditions

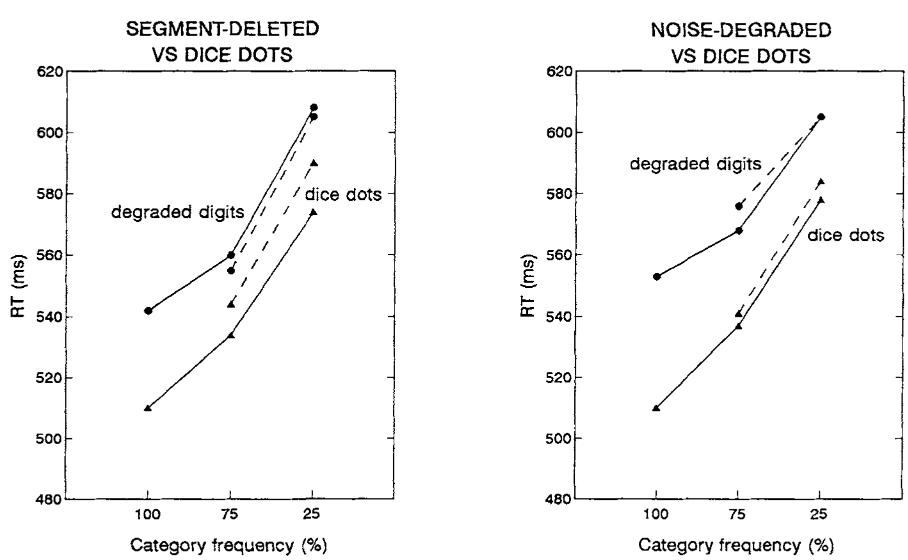

Figure 11. Observed (solid lines) and predicted (dashed lines) mean reaction times (RTs) as a function of stimulus category and category frequency in Experiment 5.

with the data actually obtained in these conditions. The contrast fell far short of significance, $F(1, 28) = 1.06$, $MSE = 1{,}185.07$, $p > .30$, indicating that the fit of the model was fairly good. In fact, as Figure 11 indicates, the only serious deviations from the model were found for dice when combined with segment-deleted digits, where the predictions exceeded the mixing costs actually found. However, even here the deviation of the data from the prediction was no more than a trend, $F(1, 14) = 3.97$, $MSE = 2{,}499.78$, $p = .07$.

Discussion

The most important result from Experiment 5 was that mixing costs for dice were smaller in the context of intact digits than in the context of degraded digits regardless of the type of degradation. This result indicates that mixing costs were not caused by differences among the digit categories as to how the dots were distributed in the central region of the stimuli. If that were the case, the mixing costs for dice should have shown an increase when mixed with segment-deleted digits, intact digits, and noise-degraded digits, respectively. Again, an account of mixing costs for dice in terms of confusion is also highly unlikely because these mixing costs were just as large when dice were combined with presumably highly similar segment-deleted digits as when dice were combined with presumably highly dissimilar noise-degraded digits. Note, finally, that the data are also not consistent with a criterion account. This is most obvious for the mixing costs for degraded digits when combined with dice. Even though degraded digits showed the longest RTs in pure blocks, their mixing costs were among the largest in Experiment 5, even exceeding 50 ms in the 25% condition. This is a particularly strong example because the frequency of occurrence of each instance of the degraded category was as large in mixed blocks as in pure blocks because of the application of the instance balancing procedure.

By contrast, the results of Experiment 5 are generally consistent with the view that mixing costs proceed from shifting among computational processes taking place on an established base representation. In the General Discussion section I elaborate this view on the basis of the observed additivity and asymmetry of mixing costs.

General Discussion

The objective of this research was to explore the nature of processing underlying perceptual mixing costs. The approach to this issue was to present several stimulus categories pairwise into the pure/mixed design and to interpret the size of the resulting mixing costs in light of the dimensions on which the categories mutually differed. This way it was possible to segregate dimensions that were relevant for the occurrence of mixing costs from those that were irrelevant. In turn, the relevant dimensions provided helpful insights into the mechanism underlying mixing costs.

By means of this approach, the present study provided considerable evidence against criterion models for perceptual mixing costs. Conceiving of the criterion as a deadline, such a model predicts mixing costs to the extent that stimulus categories show different RTs in pure blocks (e.g., Lupker et al., 1997). The present study showed, however, that a pure-block RT difference was neither a necessary nor a sufficient condition for the occurrence of perceptual mixing costs. In particular, Experiments 2 and 3 showed pronounced mixing costs for categories that did not show a clear pure-block RT difference, whereas Experiment 4 showed an absence of mixing costs for categories that showed a clear pure-block RT difference. These findings, together with the observation that no shifts in the speed–accuracy trade-off occurred, show that criterion models cannot adequately account for the present mixing costs.

One could object that these findings argue only against a deadline model (e.g., Lupker et al., 1997), not against a criterion model that assumes that the criterion is set in the information domain (e.g., Grice, 1968; Strayer & Kramer, 1994a, 1994b). Under the latter model, the finding that different stimulus categories have equal pure-block RTs does not allow the inference that their response criteria occupied equal positions in the information domain. Alternatively, a relatively slow accumulation of information over time may be compensated for by the use of a less conservative criterion, resulting in about equal RTs for categories in pure blocks, with a higher level of accuracy for the category with the faster accumulation of information. This could have been the case in the present study, where in pure blocks segment-deleted digits were more accurately responded to than noise-degraded digits, whereas their mean RTs were about equal. Still, it seems that this account cannot predict typical patterns of mixing costs. Consider two hypothetical cases, in which one category, A, has a lower rate of information accumulation over time but is responded to at a less conservative criterion in pure blocks than the other category, B. First, on the basis of pure-block RTs, Category A turns out to be the slow category because its less conservative criterion does not fully compensate for its lower accumulation rate. In this case, responding at a common criterion in mixed blocks should result in larger mixing costs for the slow Category A than for the fast Category B (cf. Grice, 1968), which is opposite to the typical asymmetry of mixing costs as presently found in Experiments 1, 2, and 5. Second, Category A turns out to be the fast category because its lower accumulation rate is overcompensated for by its less conservative criterion. This bizarre case correctly predicts the asymmetry of mixing costs, but it makes the prediction that the fast Category A becomes the slow category when responding occurs at a common criterion in mixed blocks. I do not know of any research reporting this reversal.

At least two more findings from the literature argue against a criterion mechanism for perceptual mixing costs. First, Sanocki (1988, Experiment 2) used accuracy instead of RT as a dependent measure in a task requiring the participant to identify letters occurring in briefly presented strings. In pure blocks, the type font of the letters was invariant across trials, but in mixed blocks, two highly distinct type fonts alternated on successive trials. The results showed that letter identification was less accurate in mixed blocks than in pure blocks. Because of the absence of speed stress in this task, these mixing costs cannot be interpreted in terms of a criterion model and therefore show reduced sensitivity of the perceptual system to font-specific characteristics in mixed blocks. Second, in a previous study (Los, 1997) I manipulated the response criterion directly by comparing a condition of high speed stress to a control condition of normal speed stress. I used the same task as in the present study, which required participants to identify noise-degraded and segment-deleted digits in pure, mixed-1:3, and mixed-3:1 blocks. The results showed that mixing costs were just as large under high speed stress as in the control condition. Again, this result disconfirms the prediction of a criterion model that mixing costs critically depend on the position of the criterion, as has been shown by Strayer and Kramer (1994b, Experiment 5).

Instead, the results of the present study suggest that perceptual mixing costs occur whenever stimulus categories require different computational processing. This view is not readily supported because stimulus categories that seem to require different computational processing typically also differ on dimensions that seem trivial from a computational point of view but may still contribute to mixing costs. The present study aimed at isolating several of these dimensions so as to assess their contribution to mixing costs. Thus, it was possible to show that there was no contribution to mixing costs from dimensions such as the form of the frame surrounding the digit (Experiment 4), the distribution of dots across the stimulus area (Experiment 5), and global perceptual similarity (Experiment 5). These findings, then, lend credibility to an alternate processing view. In the final part of this article, I elaborate this view.

Mixing Costs and the Nature of Processing

Ullman (1984) envisaged a visual routine as an assembly of elemental visual operations. Within the assembly, the elemental operations work together in a coordinate way as they are applied to a base representation. The base representation results from early visual processing in a fully bottom-up way, that is, in a way completely determined by the input. Visual routines are top-down applied to selected parts of the base representation, so as to obtain information on objects and relations among objects for further analysis.

In the introduction to Experiment 5, I proposed that perceptual mixing costs may reflect shifting among visual routines. This proposal could be vindicated from a computational point of view by an analysis of what it takes to identify a stimulus belonging to a certain category. Thus, there is good reason to assume that digits and dice dots, the major stimulus categories used in Experiment 5, call on different visual routines. Ullman (1984; see also Trick & Pylyshyn, 1994) proposed a visual routine for the enumeration of elements (e.g., dice dots) composed of two elemental operations: "indexing," the process of shifting the processing focus to selected locations in the base representation, and "marking," the process of indicating those locations in the base representation where the processing focus has been before. In its simplest form, enumeration could then be described as a process of initializing a counter at zero, followed by the cycle of indexing an element, increasing the counter by one, and marking the element, until all elements are marked. For the identification of intact digits, the elemental operations of indexing and "boundary tracing" could make up a visual routine (Ullman, 1984). Boundary tracing is the process of tracking a contour in the base representation. Ullman (1984, p. 144) suggested that this process enables the decomposition of a digit at points of maxima in curvature. The resulting constituent strokes may then enable identification. In a similar vein, different visual routines may be proposed for the identification of intact, noise-degraded, and segment-deleted digits. It seems reasonable to assume that boundary tracing plays an important role in the identification of all these types of digits. But whereas this operation may directly enable the detection of the points of maxima in curvature for intact digits, for degraded digits it may be necessary first to infer the correct contours from among several possibilities. For this purpose, additional clean-up and fill-in operations should be postulated. This analysis, then, suggests that perceptual mixing costs are caused by switching among visual routines.[5]

A major challenge to this view comes from the additivity of perceptual mixing costs found in Experiment 5. This additivity is particularly interesting because it was not found by Allport et al. (1994) in the realm of task switching. A straightforward interpretation of additive mixing costs is that switching takes place during separate stages, analogous to the interpretation of additivity in the additive factors method (Sternberg, 1969). Indeed, if switching among visual routines occurs during a single stage, it is not obvious why the mixing costs in the dice/degraded context would be the sum of the mixing costs in the intact/degraded and dice/intact contexts.

An alternative view may therefore be preferred in which switching among processes called on by digits of different perceptual categories occurs at a different stage of processing than switching among processes called on by digits and dice. For instance, the mixing costs in the dice/intact digits context may occur at the level of visual routines, as assumed before, but the mixing costs for different types of digits may occur at an earlier stage. In this respect it is interesting to consider the principle of fine-tuning of processing parameters, as embodied in Sanocki's (1987, 1988, 1991) descriptions model. With it, Sanocki (1988) accounted for the mixing costs he observed when participants identified letters

[5] The feasibility of this view is called into question by the paradox that mixing costs are exogenous in origin (as pointed out in the introduction and as moderately supported by Experiments 1 and 2), whereas visual routines are applied to the base representation in a top-down way. In fact, in a previous study (Los, 1997) I found that participants cannot reduce their mixing costs when they are informed at the start of a trial of the forthcoming stimulus category, which they should be capable of doing if mixing costs are mediated by visual routines that are under complete top-down control. The paradox may be resolved, however, by proposing that the top-down selection of a visual routine is possible only after a stimulus has been presented, and not in the absence of stimulus material. This seems a viable possibility, even though it implies a limitation of the top-down applicability of visual routines.

of different font types that were presented in the pure/mixed design. A font type can be characterized by certain abstract regularities, such as the ratio of letter bodies and their terminating lines, which are shared by letters of the same font type but are not likely to be shared by letters of different font types. The descriptions model makes use of these regularities by saving font-specific parameters across different string presentations. These parameters enable an efficient translation from an actual letter instance to a "deep" letter representation. If the font type of letter strings shifts from one presentation to the next, the font-specific parameters may fail to carry out the translation and require a time-consuming adjustment. In a similar vein, it could be that switching among digits of a different type of degradation also requires time-consuming parameter adjustments. When these adjustments occur at a different stage than switching among visual routines, their entailing switch costs may well add to those of switching among visual routines.

From this discussion it should be clear that the precise nature of the processing required for the identification of the various stimulus categories cannot be established on the basis of the present results. The additivity of mixing costs is intriguing, though, and may provide a valuable tool for future research once its interpretation derives from a firmly established theory.

Mixing Costs and the Nature of Switching

I have collected evidence that perceptual mixing costs are attributable to switching among perceptual processes, but I have left untouched the important question of how these processes interact. I would like to conclude by suggesting two principles that may govern this interaction. The first principle is that within a given processing stage, the activation of one process implies the suppression of other processes. One way this could be achieved is by assuming that the processes are organized in a competitive framework, realized by inhibitory mutual connections. So, a switch of processing involves the activation of processes that have been inhibited shortly before. In line with this principle is the observation that mixing costs are typically not evenly distributed across the trials of mixed blocks but are more pronounced on alternation trials (as observed in Experiments 1 and 2; see also Los, 1997; Maljkovic & Nakayama, 1994). The second principle holds that the period during which a process is active determines the amount of inhibition transferred to other processes of the same stage. That is, the longer a process operates, the greater the suppression of other processes. This principle provides the basis for explaining the asymmetry of mixing costs—namely, that mixing costs are larger for the fast level of a variable than for the slow level (as observed in Experiments 1, 2, and 5; see also Los, 1996). The asymmetry follows directly from the second principle because processes analyzing the slow level of a variable are assumed to be active during a longer period than are processes analyzing the fast level. Thus, the system must overcome more inhibition when switching from the slow level to the fast level than vice versa, thereby incurring stronger mixing costs when analyzing the fast level.[6]

There are at least two good reasons to prefer a mechanism based on these two principles over an apparently simpler mechanism assuming that a repeated routine takes advantage of residual activation left over from the previous trial. First, a mechanism emphasizing the virtue of residual activation has difficulties explaining that mixing costs are generally larger for the category that is responded to faster in pure blocks. In fact, such a mechanism would rather predict the opposite result, because the presumably elaborate computations required for the slow category leave more room for improvement than the simpler computations required for the fast category. Second, a mechanism emphasizing the virtue of residual activation would qualify the term *mixing costs* as being inferior to the alternative term *blocking benefits*. In a previous study (Los, 1994), I addressed this issue by relating the effects of mixing costs to a baseline. In the experimental condition of that study, participants identified an intact or degraded visual digit (S2) briefly after they identified another intact or degraded visual digit (S1). In the baseline condition, participants identified the same intact or degraded S2 as in the experimental condition but did so briefly after they identified the pitch of a nondegraded pure tone. Thus, on trials in the baseline condition, participants did not engage in visual processing before S2 was presented. The dependent variable of interest was RT2, the time taken to identify S2. Relative to the baseline, the results of the experimental condition showed one clear deviation: RT2 was relatively long when an intact S2 was preceded by a degraded S1. Thus, that study yielded powerful evidence for mixing *costs* for the fast intact category and thereby clearly favors a mechanism based on the above two principles over a mechanism emphasizing the virtue of residual activation.

More generally, it may be noted that the functional significance of the inhibitory control proposed here is to stabilize ongoing activity by preventing other activity from interfering. This is not a new function, because it is embedded in several control mechanisms proposed for task switching (Allport et al., 1994), access to working memory (Hasher & Zacks, 1988), and interference in negative priming paradigms (Driver & Tipper, 1989; Tipper, 1985).

---

[6] This interpretation of asymmetric mixing costs may provide an additional argument against the idea that switching in the context of intact and degraded digits takes place at the level of visual routines. As was suggested in the previous section, the identification of an intact digit can be exhaustively described by a subset of elemental operations involved in the identification of degraded digits. In that case, though, it is not clear how a mechanism of lateral inhibition can account for the mixing costs for intact digits without additional assumptions because not one elemental operation involved in the identification of intact digits is suppressed during the processing of a degraded digit. I am grateful to Steven Yantis for bringing this point to my attention.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious* *and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Baddeley, A. (1986). *Working memory.* Cambridge, MA: MIT Press.

Beringer, J. (1992). *Entwerf einer Anwendersprache zur Steuerung psychologischer Reaktionszeitexperimente* [Design of an application language for conducting psychological reaction time experiments]. Frankfurt am Main, Germany: Lang.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13,* 90–102.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203,* 237–260.

Campbell, K. C., & Proctor, R. W. (1993). Repetition effects with categorizable stimulus and response sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1345–1362.

Colombo, L., & Tabossi, P. (1992). Strategies and stress assignment: Evidence from a shallow orthography. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 319–340). Amsterdam: Elsevier.

Driver, J., & Tipper, S. P. (1989). On the nonselectivity of "selective" seeing: Contrasts between interference and priming in selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 304–314.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: II. Cognitive processes and performance* (pp. 34-1–34-41). New York: Wiley.

Grice, G. R. (1968). Stimulus intensity and response evocation. *Psychological Review, 75,* 359–373.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Ilan, A. B., & Miller, J. O. (1994). A violation of pure insertion: Mental rotation and choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 520–536.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41,* 291–297.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14,* Whole No. 89, p. 81.

Kirby, N. H. (1980). Sequential effects in choice reaction time. In A. T. Welford (Eds.), *Reaction times* (pp. 129–172). London: Academic Press.

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. In S. Kornblum (Eds.), *Attention and performance IV* (pp. 259–288). New York: Academic Press.

Los, S. A. (1994). Procedural differences in processing intact and degraded stimuli. *Memory & Cognition, 22,* 145–156.

Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica, 94,* 145–188.

Los, S. A. (1997). *Identifying stimuli of different perceptual categories in pure and mixed blocks trials: Evidence for stimulus-driven switch cost.* Manuscript submitted for publication.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 570–590.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22,* 657–672.

Marr, D. (1982). *Vision.* New York: Freeman.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1–20.

Monsell, S., Patterson, K. E., Graham, A., Huges, C. H., & Milroy, R. (1992). Lexical and sublexical translation from spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 452–467.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In G. E. Schwartz, R. J. Davidson, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenum.

Ollman, R. T. (1966). Fast guesses in choice reaction time. *Psychonomic Science, 6,* 155–156.

Ollman, R. T., & Billington, M. J. (1972). The deadline model for simple reaction times. *Cognitive Psychology, 3,* 311–336.

Paap, K. R., Williams, R. W., & Johansen, L. S. (1992). Dual-route models of print to sound: Red herrings and real horses. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 293–318). Amsterdam: Elsevier.

Pashler, H., & Baylis, G. (1991). Procedural learning: 2. Intertrial repetition effects in speeded choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 33–48.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica, 74,* 123–167.

Sanocki, T. (1987). Visual knowledge underlying letter perception: Font-specific, schematic tuning. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 267–278.

Sanocki, T. (1988). Font regularity constraints on the process of letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 472–480.

Sanocki, T. (1991). Intra- and interpattern relations in letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 924–941.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 598–616.

Sokolov, E. N. (1963). *Perception and the conditioned reflex.* Oxford, England: Pergamon Press.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89,* 669–679.

Sperling, G., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: I. Sensory processes and perception* (pp. 2-1–2-65). New York: Wiley.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Erlbaum.

Stoffels, E. J. (1996). Uncertainty and processing routes in the selection of a response: An S-R compatibility study. *Acta Psychologica, 94,* 227–252.

Strayer, D. L., & Kramer, A. F. (1994a). Strategies and automaticity: I. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 318–341.

Strayer, D. L., & Kramer, A. F. (1994b). Strategies and automaticity: II. Dynamic aspects of strategy adjustment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 342–365.

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 89–103.

Tipper, S. P. (1985). The negative priming effect: Inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 37A,* 591–611.

Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review, 91,* 68–111.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102.

Ullman, S. (1984). Visual routines. *Cognition, 18,* 97–159.

Van Duren, L. L., & Sanders, A. F. (1988). On the robustness of the additive factors stage structure in blocked and mixed choice reaction designs. *Acta Psychologica, 69,* 83–94.

Received November 23, 1993
Revision received September 23, 1997
Accepted October 27, 1997 ■